IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 December 2022, accepted 20 January 2023, date of publication 10 February 2023, date of current version 16 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3244088

TOPICAL REVIEW

Datasets for the Quality Assessment of Light Field Imaging: Comparison and Future Directions

EDRIS SHAFIEE, (Graduate Student Member, IEEE), AND MARIA G. MARTINI, (Senior Member, IEEE)
Wireless and Multimedia Networking Research Group, Kingston University London, KT1 2EE London, U.K.
Corresponding author: Edris Shafiee (e.shafiee@kingston.ac.uk)

ABSTRACT With the increasing research focus on light field imaging in recent years, it has become essential for researchers in this field to either benefit from access to equipped laboratories with light field acquisition devices and displays or to have access to publicly available light field imaging datasets. Some datasets are indeed available, each with a different nature. For instance, some contain real world images or sources while others are based on purely synthetic images or sources generated by computer graphic tools; others are a combination of both. Datasets for the quality assessment of light field content include pristine light field content as well as sources affected by different levels of impairments. The latter are tested subjectively by a panel of viewers and often objective metrics are also calculated. This paper presents a comprehensive comparative review of 33 publicly available datasets that span from content-only datasets to specific task based datasets and quality assessment datasets. While our aim is to review and investigate what each dataset has to offer and which tests had been considered by their proposers, we also take the opportunity to leverage the results of previous studies to identify and discuss the challenges ahead and identify the areas with potential for improvement.

INDEX TERMS Light field displays, quality assessment, dataset, subjective tests, objective quality metrics.

I. INTRODUCTION

L ight field [1], also known as integral or plenoptic imaging, is in recent years at the forefront of research attention, in particular for the the ability to provide the viewer with an enhanced sense of immersion in the scene. As opposed to conventional two-dimensional (2D) images, which record and contain the 2D projection of the scene light rays, a light field contains the light rays in space across all directions, hence much richer data from the scene than a 2D image. The concept of distribution of light field was first introduced in 1939 by Greshun [2]. In 1991, Adelson et al. [3] completed the relevant model, describing what is known since then as plenoptic function. The plenoptic function is a multidimentional function that takes into account the location (x, y, z) in which the light ray is obtained (camera position), the angles (θ, ϕ) of the light ray (azimuth and orientation), the wavelength λ of the light ray (representing color) and finally the time t when the light ray was captured. Therefore this is a 7D function, denoted as $L(x, y, z, \theta, \phi, \lambda, t)$. However, computationally handling a 7D function is extremely difficult and, as mentioned by Wu et al. [4], the 7D function can be simplified to 5D by assuming the plenoptic function to be monochromatic and time invariant in order to decrease the complexity and volume of the data consumed (three reduced plenoptic functions can be considered - one for each RGB colour component - for colour content and temporal information can be captured via different frames for dynamic content). Levoy and Hanrahan [1] simplified this function by realising that the radiance of the light ray in free space remains constant along a straight line, hence reducing the total needed dimensions to four. For instance, a 4D light field can be represented via the coordinates of the intersections of a light ray with two planes placed at arbitrary positions, e.g. (u, v) and (r, s), leading to a plenoptic function $L(u, v, r, s)$. Based on this representation, the light field can be represented

The associate editor coordinating the review of this manuscript and approving it for publication was Alba Amato.

as a 2D array of images captured by cameras defined by the two parallel planes (u, v) and (r, s).

The capability of light fields to capture a rich amount of spatial-angular information of an object or a scene at once, and to allow researchers to work on the data post-acquisition, made this research area highly attractive. Such technology has not yet widely reached the consumer market, and the role of researchers in this field is crucial to tackle the obstacles preventing this technology to reach the end users. In order to support the research community in the area, it is important to have available datasets that can be used by researchers to test and compare their ideas and methods, including compression and transmission strategies [5] and objective quality assessment methods (discussed in Section II-C). Thanks to the efforts of several researchers, a number of light field datasets have in the past years been made publicly available.

In this paper we provide a detailed description and a taxonomy of existing publicly available light field datasets, with particular focus on public datasets for the quality assessment of light field imaging. We also discuss to what extent the different datasets enable a comparison of objective and subjective quality metrics for the different datasets, highlighting which quality metrics have been designed and tested on each dataset. We finally focus on the limitations of existing datasets and provide some recommendations for the design of new datasets that can support the scientific community in the design of novel compression, transmission, and quality assessment algorithms for light field content. Some earlier papers, providing results for objective quality metrics or specific tasks performed on light fields, listed some datasets; however, the purpose was not to describe and compare them and only a limited number of datasets was considered in each of these papers [6], [7], [8], [9], [10]. A review of some of the existing datasets was provided a few years ago by Ellahi et al. [11]. The paper listed and described some datasets that were public at the time of publication. With the fast evolution in this research area, many datasets have been published recently. In our paper we have considered a much larger number of datasets in the effort to provide an exhaustive review and we have dived more in depth in the details and use cases of such datasets, covering content features, description and comparison of the deployment of the subjective tests, including demographic information of subjects participants and the results of objective and subjective assessment on those datasets. We believe this work can support research scientists working in the areas of light field compression, transmission, and quality evaluation in the selection of the most appropriate datasets for developing and testing their algorithms, as well as facilitate comparison among the works of different research teams worldwide. We also hope that this paper will stimulate researchers to address the lacks of existing datasets and provide public datasets useful to answer the most recent research questions. The ongoing development of technical standards in the area of light field will also benefit from this work. Part of this review work has been performed to support the development of an international IEEE standard (IEEE P.3333.1.4).

The remainder of this paper is structured as follows. Section II reviews all the datasets that are selected in this paper and provides a relevant taxonomy. In Section III we discuss the current challenges in generating and using datasets and different aspects of objective and subjective assessment. Section IV concludes the paper and our review.

II. LIGHT FIELD DATASETS

Table 1 lists the light field datasets that are discussed in this paper. We divide the reviewed datasets in three different groups: content only datasets, where only Light Field (LF) content is provided, task based datasets, i.e., datasets where content is provided together with information on specific tasks (e.g., disparity or depth estimation, material recognition, scene reconstruction), and finally Quality of Experience (QoE) datasets, where content and information on the relevant quality assessment is provided, e.g., Mean Opinion Score (MOS). Being publicly available, such datasets can be used in research works for different purposes, but we report in the table the initial purpose behind their creation. Table 1 also illustrates the acquisition device and the number of sources and images in each dataset and whether the datasets contain real world images, synthetic images or a mix of both. The datasets are listed in chronological order and the table also reports the first author/creator of each dataset.

Figure 1 illustrates the taxonomy of the datasets. The same categories mentioned above for Table 1 are applied in the taxonomy with further categorisation on acquisition method and, for QoE datasets, type of display used for subjective tests. Some datasets are listed in different acquisition method sub groups, as they contain content acquired with different methods.

We discuss the details for each of the identified groups in the following.

A. CONTENT-ONLY LF DATASETS

These are datasets where only the content is published, without results of objective or subjective quality assessment tests or other data on task performance. 15 datasets of this type are listed in Table 1. Due to space limitations, in this section we review datasets [12], [13], [14], [15], [18], [19], and [20].

The (New) Stanford Light Field Archive [12] provides a database of four light field datasets acquired by different acquisition devices listed in the first row of Table 1. This database provides a rich source of light fields from different real world LF SRC and angular views. The four different acquisition methods provide a diverse range of the number of views for each object with different spatial resolution. The four datasets generated with different acquisition methods, as illustrated in Figure 2, are as follows:

(i) Light Fields from Lego Gantry: This dataset contains 13 sources acquired with a Canon Digital Rebel XTi mounted on a Lego Mindstorms gantry. For each source

TABLE 1. Light field datasets.

Dataset	Contributors	Year	Dataset Class	Acquisition Device	LF Source (SRC)	LF Processed Video Sequence (PVS)	LF Content	Static Objects or Scenes (S) / Video (V)
The (New) Stanford Light Field Archive [12]	Vaish et al.	2008	Content-Only	Stanford Multi-Camera Array [100 VGA video cameras] Light Field Gantry [Digital Camera] Lego Mindstorms Gantry [Canon Digital Rebel Xti] Light Field Microscope	22	5,524	Real Scenes / Objects Microscopic LF Images	S/V
EPFL Light-Field Image Dataset [13]	Rerabek et al.	2015	Content-Only	Lytro Illum	10	119	Real Scenes	S
The new light field dataset [14]	Rerabek et al.	2016	Content-Only	Lytro Illum	10	118	Real Scenes	S
Synthetic Light Field Archive [15]	Wetzstein [16]	2013	Content-Only	Synthetic	18	Many Images in each dataset	Artificial static and animations	S
Ziegler et al. [17] (Fraunhofer IIS)	Ziegler at al.	2017	Content-Only	Sony alpha 7RII	9	9	Natural Scenes	S
R8 Raytrix Camera VIDEO Dataset [18]	Guillo et al.	2018	Content-Only	R8 Raytrix	3	900	Real Objects	V
Turntable General Dataset [19]	Tamboli et al.	2018	Content-Only	Nikon D5300 DSLR	7	15,120	Real Objects	S
Comprehensive light field image dataset [20]	Sheng Liu	2018	Content-Only	Lytro Illum	31	4,251	Many scenes with different lighting and environmental conditions	S
Light-Field Intrinsic Dataset [21]	Shekar et al.	2018	Content-Only	Canon EOS 6D/5D Sony Alpha 7 RII	12	12	Synthetic Real Objects	S
The Stanford Multiview Light Field Datasets [22]	Dansereau et al. [23]	2019	Content-Only	Lytro Illum	994	7,253	Multiview Datasets: 30 Categories Three View Datasets: 421 Indoor/ 543 Outdoor Scenes	S
Non-planar inside-out dense light field dataset [24] (Fraunhofer IIS)	Zakeri et al.	2019	Content-Only	Sony alpha 7RII	2	2	Non-planar Light field Representation	S
TEDDY [25] (Fraunhofer IIS)	Gul et al. [26]	2020	Content-Only	Sony alpha 7RII	6	6	Real Objects	S
CIVIT Dataset [27] (Tampere)	Moreschini et al. [28]	2020	Content-Only	Optronis CP70-12-C-188 with Nikon Fx 35 mm f/1.8 lens	6	1,358	Synthetic Real Objects	S
Yue et al. [29] (Cambridge Fraunhofer IIS)	Yue et al. [30]	2020	Content-Only	Sony Alpha a7II	15	15	11 Synthetic scenes 4 Real Scenes	S
4DLFVD: A 4D Light Field Video Dataset [31]	Hu et al. [32]	2021	Content-Only	100 monocular video cameras [10 × 10]	9	9	Real Scenes	V
Disney High Spatio-Angular Resolution Light Fields [33]	Kim et al.	2013	Scene Reconstruction	Canon EOS 5D Mark II	5	555	Real Indoor / Outdoor Scenes	S
Densely Sampled 4D Light Fields Dataset [34]	Wanner et al	2013	Disparity Estimation	Synthetic and Nikon D800	13	13	7 Synthetic / 6 Real Objects	S
Lytro Dataset [35]	Mousnier et al.	2015	Light Field Reconstruction	Lytro Illum	30	30	Real Scenes/Images	S
The 4D Light-Field Dataset of Material [36]	Wang et al.	2016	Material Recognition	Lytro Illum	12	1,200	Real Objects	S
4D Light Field Dataset [37]	Honauer et al.	2017	Depth Estimation	Synthetic using Blender	24	24	Synthetic	S
Local Light Field Synthesis [38]	Srinivasan et al. [39]	2017	synthesizing a 4D RGBD light field	Lytro Illum	Several Scenes	3,343	Several Plant based scenes such as roses, poppies, thistles orchids, lilies, irises	S
Multi-View Light-Field Video [40]	Sabater et al. [41]	2017	Depth-based rendering algorithm for Dynamic Perspective Rendering	16 IDS Cameras with CMOSIS (CMV2000) sensors	12	12	Real Scenes	V
Matching light field datasets from plenoptic cameras 1.0 and 2.0 [42]	Ahmad et al. [43]	2018	Comparison of Lytro Illum and Raytrix R29	Lytro Illum Raytrix R29	31	31	Real Objects	S
The Plenoptic 2.0 Toolbox [44]	Palmieri et al. [45]	2018	Depth Estimation	Raytrix R29 and R42 Synthetic	19	19	13 Real Objects 6 Synthetic	S
Multispectral Light Field Dataset [46]	Schambach et al. [47]	2020	Light Field Deep Learning	Synthetic	507	507	Synthetic Scenes	S
SMART [48]	Battisti et al.	2016	QoE	Lytro Illum	16	16	Real Scenes	S
Turntable Quality Dataset [49]	Tamboli et al.	2016	QoE	Basler's ACE 1300gc	3	3	Real Objects	S
MPI-Light Field [50]	Adhikarla et al.	2017	QoE	Canon EOS 5D Mark II	14	350	9 Synthetis Scenes 5 Real Scenes	S
VALID [51]	Viola et al.	2018	QoE	Lytro Illum	5	5	Real Scenes from [14]	S
Win5-LID [52]	Shi et al.	2018	QoE	Lytro Illum	10	220	6 Real Scenes from [14] 4 Synthetic Scenes from [37]	S
Shan et al. [53]	Shan et al.	2018	QoE	Lytro Illum	6	6	Real Scenes from [13]	S
LFDD: Annotated LF Display [54]	Zizen et al.	2020	QoE	Synthetic	10	10	Synthetic Scenes	S
V-SENSE [55]	Gill et al.	2020	QoE	Synthetic and Lytro Illum	20	20	Real Scenes from [12], [14], [33] and Synthetic Images from [37]	S

there are 289 PVS from a 17 × 17 camera grid. The image resolution is different across different sources.

(ii) Microscope Light Fields: This dataset contains three sources acquired with an optical microscope [56] with a lenslet array. There are different numbers of views for each source as the grid size varied [20 × 20 grid, 16 × 16 grid]. The light fields in this dataset are of low resolution as they were captured with a single snapshot from a single camera, hence not many samples were captured to record all the spatial and angular information. Therefore, because of the low spatial resolution, they look pixelated.

(iii) Light Fields from the Gantry: four sources were acquired using a digital camera and a computer-controlled gantry. The number of views were different in this acquisition method as the grid size varied [21 × 5 grid, 16×16 grid]. The spatial resolution of the images in the dataset was also different across different sources. The sources in this dataset offer good opportunities for research in refocussing and three-dimensional (3D) reconstruction, whereas the Lego gantry dataset sources offer a wider range of research opportunities owing to the complex and highly detailed sources.

(iv) Light Fields from the Camera Array: two sources with 88 and 45 views with image resolution of 640 × 480 were acquired using 100 VGA video cameras [57] controlled by four computers.

Rerabek and Ebrahimi [14] proposed "The New Light Field Dataset". This dataset was first proposed by the authors in Light Field Raw (LFR) file format as provided by the

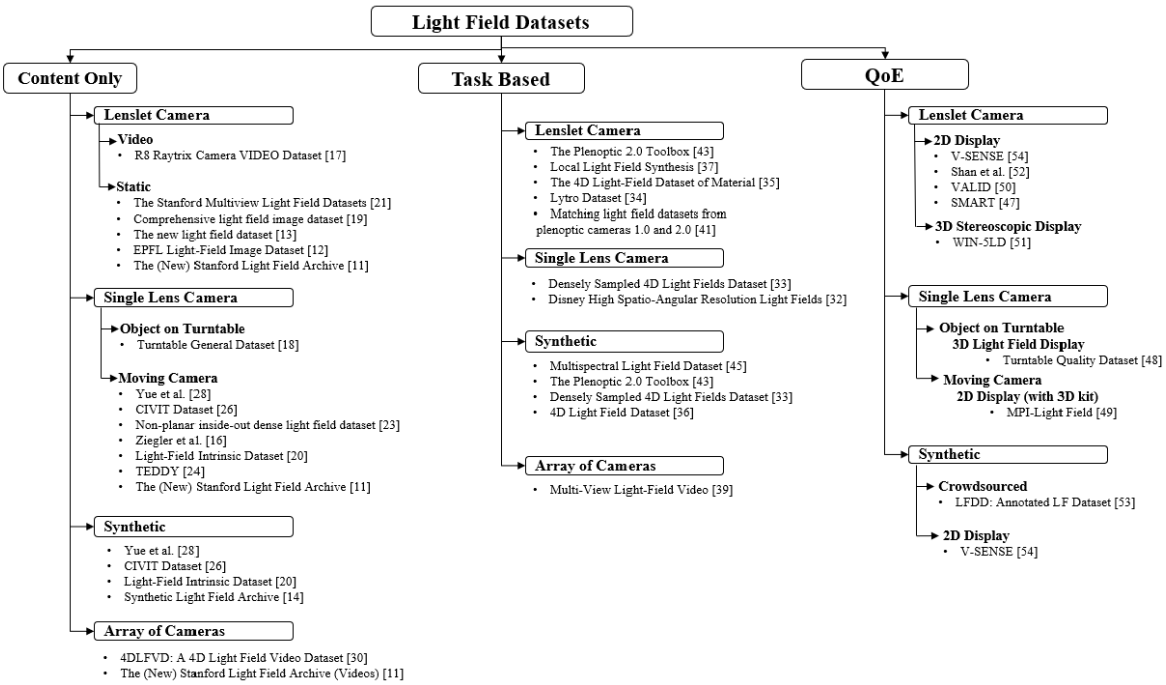
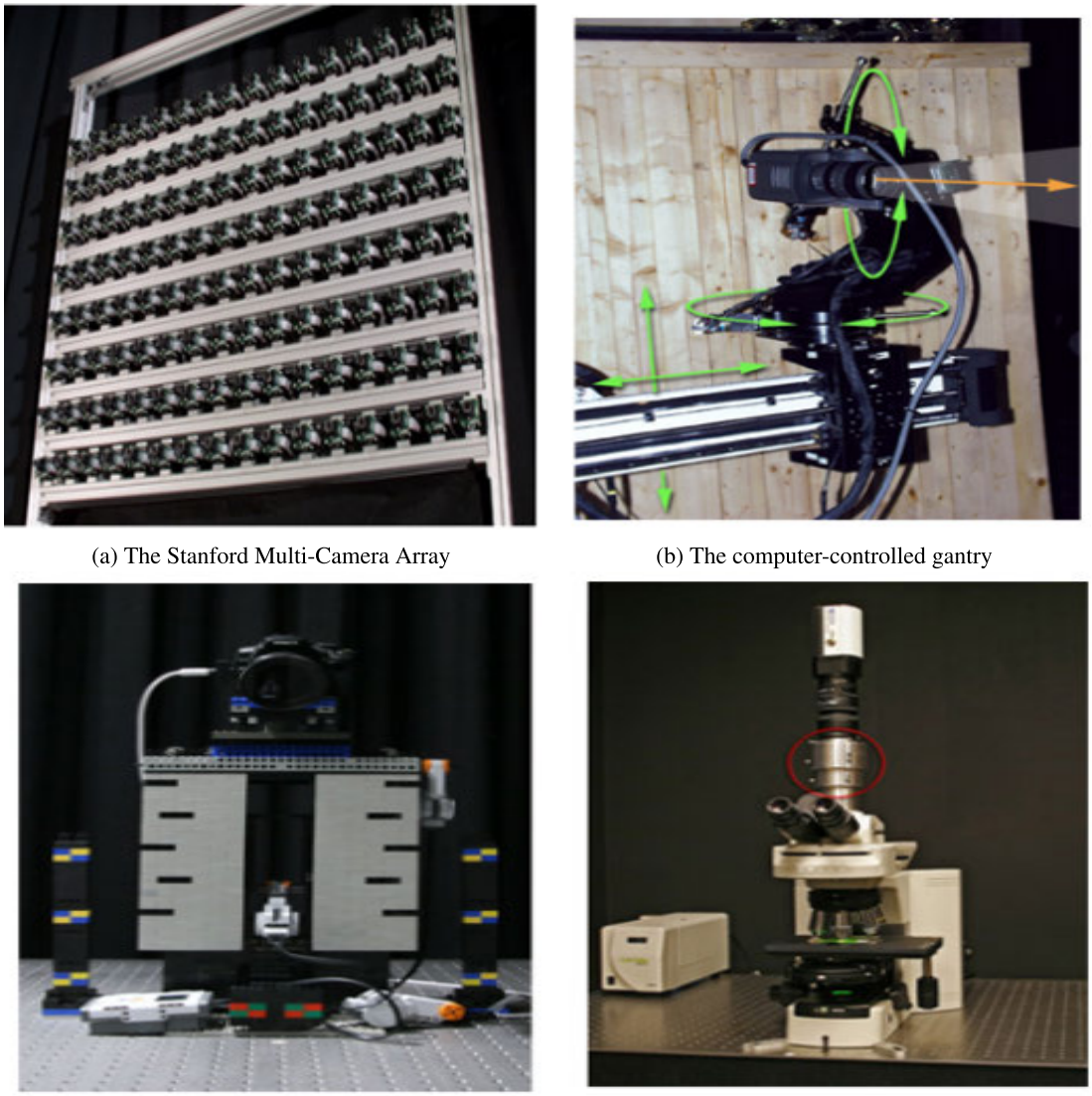


FIGURE 1. Light field datasets taxonomy.

Lytro Illum camera in 2015 with the name ''EPFL Light-Field Image Dataset'' [13]. The ''EPFL Light-Field Image Dataset''[1] consisted of 119 LFR images from 10 different real world objects/sources with resolution of 5368 × 7728 pixels. As well as the LFR file format images, ''The New Light Field Dataset'' consists of 4D light field representation of the images that were extracted from original LFRs using the Light Field Matlab Toolbox[2] [58]. In addition, the dataset contains thumbnails, the corresponding depth maps, the relevant depth of fields coordinates, and camera calibrating data. The four-dimensional (4D) LF images that were generated by the Light Field Matlab Toolbox have 15 × 15 views 434 × 625 resolution of each view. 4D corresponds to RGB and the additional weighting image component. The wide range of real world sources in this dataset makes it suitable for further research in objective and subjective assessments. As will be noted in the Quality Assessment Dataset section, the VALID [51] and WIN5-LID [52] datasets used for their subjective quality assessment studies six images from this dataset.

Tamboli et al. proposed a Turntable General Dataset[3] [19] using an automated Photobench 360 turntable[4] and a Nikon D5300 DSLR camera.[5] Seven distinct objects were positioned on the turntable for this dataset. The camera was manually moved over a slider between three positions, 2 cm apart from each other, to capture images from three different views from the target object at 0.5 degree of turntable revolution, which was achieved by a software controlling the turntable stop/start intervals, resulting in a total 720 images per camera per object. Therefore the angular resolution of the proposed dataset is 0.5 degrees [two views per one degree] and a total of 15,120 images of 6000 × 4000 pixels resolution were generated. This dataset offers 2,160 three-view stereo images per object which can be used for research in view interpolation techniques. To ensure stability of the objects during the capturing phase, the base of the object was placed on the turntable firm enough to avoid any movement on table rotation. As the same fixed camera was used to capture images at different angles, the inter-camera variations were eliminated throughout the acquisition phase. Also three view stereo enables robust depth estimation. To ensure turntable consistency, an image from one camera position and its corresponding image from the adjacent camera position were obtained and three quality metrics were evaluated on the pair of images. In summary, this dataset appears to be the first proposed high angular resolution [2 views per degree] dataset using a turntable. For every angular position - 0.5 degree - there are three views separated by known distance from each other. The angular resolution of two views per one degree of this dataset satisfies the findings of Kara et al. [59] that suggested that, in order to have a good glasses free end user experience of Light Field visualisation, light field content should have an angular resolution of at least 1 degree. This angular resolution also satisfies another finding of

[1] http://mmspg.epfl.ch/EPFL-light-field-image-dataset
[2] https://www.mathworks.com/
[3] https://www.iith.ac.in/DReaM/
[4] https://www.ortery.com/photography-equipment/photocaptureseries/photocapture-360/
[5] http://imaging.nikon.com/lineup/dslr/d5300/



(a) The Stanford Multi-Camera Array

(b) The computer-controlled gantry

(c) Lego Mindstorms gantry

(d) The light field microscope

FIGURE 2. The (new) stanford light field datasets [12] acquisition devices.

Kara et al. [60], [61] from separate experiments that, in order to give the non stationary viewers the opportunity of experiencing continuous horizontal parallax, an angular resolution of less than 1 degree was required.

Guillo et al. [18] proposed a light field video dataset using an R8 Raytrix video camera. This dataset is the only dataset in our comparative investigation that was generated using such a camera. For each light field video sequences of the three sources in this dataset, the camera's specific software (Raytrix API[6]) was used to extract 25 views per frame from the light field. The 2D spatial resolution of each frame is 1920×1080. The first two sources were created with objects on a turntable and the camera is fixed. The third source is static and the camera revolves around the object. All three video sequences are captured at 30fps. A disparity map is also computed for each view of each frame.

Sheng Liu [20] proposed the ''Comprehensive light field image dataset''. This publicly available large light field image dataset consists of 4,251 light field images in 31 different categories that were captured using a Lytro Illum light field camera. The LFR images were with resolution of 5250×7574 pixels. As a result of Lytro Illum camera's 14×14 lens array, the 2D spatial resolution of the processed images becomes 375×541. Within each category, many sources were considered, with different lighting and environmental conditions in the capturing process This dataset contains static as well as dynamic objects/sources images. The depth map of the images was also investigated to determine the depth of the objects in a source and determine if an object was farther away or closer to camera than the camera's focal point.

[6]https://www.raytrix.de/Rx.ApiLF.3.1/index.html

FIGURE 3. Dragon and bunnies scene from synthetic light field archive [15]. Features: motion parallax, translucency, reflections, specularity, high-frequency out-of-focus objects, and occlusions.

This is a rich dataset in terms of the wide range of objects the quantity of the images available for further research.

Wetzstein [16] provided an archive of 18 synthetically generated static images and animation datasets [15]. The light fields in these datasets are rendered as Portable Network Graphics (PNG) images with different views of 3D sources. Each dataset in this archive has images or animations with specific features such as motion parallax, translucency, reflections, specularity, high frequency out-of-focus objects, occlusions and reflections. Figure 3 is extracted from the Dragon and Bunnies dataset [15] from this archive.

B. TASK BASED DATASETS

In the Task Based Datasets group we have selected 10 datasets which were proposed with an initial objective experiment from the proposers. Datasets [34], [35], [36], [37], and [62] are being reviewed in this paper.

Wanner et al. [34] proposed the "Densely Sampled 4D Light Fields Dataset"[7] in 2013 with the aim of evaluating disparity estimation algorithms. The 13 high quality densely sampled light field images in this dataset are divided into two categories: the Blender and the Gantry. The Blender category consists of 7 sources rendered using the open source software Blender [63] resulting in 7 synthetic/artificial images with complete ground truth disparity for all views. Four light field images of these datasets also have ground truth segmentation. The Gantry category provides 6 real-world 4D light fields captured with a Nikon D800 camera mounted on a gantry device. The angular resolution of this dataset is 9×9 and the 2D spatial resolution of the images is not the same across all images. For example, in the Gantry Category, 4 Images have resolution 898×898 pixels and 2 Images have resolution 926×926 pixels. Similarly, in the Blender category, 5 images have resolution of 768×768 pixels and the remaining two have resolution 576×1024 and 720×1024 pixels. For the Gantry category, objects were pre-scanned with a structured light scanner and were then sampled to provide ground truth ranges for the center view. The dataset contains a transparent surface with ground truth disparity for both the surface and the object behind it. As mentioned, upon creating the datasets, the aim of the authors was to evaluate the disparity algorithms performance, but this is a content dataset that can be used for other purposes.

The second dataset in this section is the "4D Light Field Dataset" proposed by Honauer et al. [37]. This dataset [64] contains 24 purely synthetic, densely sampled 4D light fields created with Blender [63] with highly accurate disparity ground truth (4 stratified created by internal renderer of Blender and 20 photorealistic light field sources created by Cycles renderer). The dataset, which was proposed to evaluate depth estimation algorithms, is divided in two groups. The first group contains 12 light fields, presented as light field benchmark. This group contains 4 stratified sources, 4 test sources and 4 training sources. The second group of 12 light fields with ground truth was presented for algorithm development and evaluation. The intention of creating a set of stratified sources was to avoid designing an algorithm that did not take into account all real world challenges, i.e., complex sources and multiple objects in a scene. The 2D spatial resolution of the 8 bit light fields created via Blender is 512×512 with 9×9 angular resolution. The dataset also contains camera parameters and disparity ranges. For the stratified and training sources, evaluation masks and 16 bit ground truth disparity maps in two resolutions (512×512 and 5120×5120) were also provided. Each scene in the benchmark group was used, depending upon its design, to assess certain factors, such as algorithm performance on the objects, assessing the effect of camera noise on the reconstruction of objects and investigating the effect of fine structures, occlusion boundaries and disparity differences and the influence of texture and contrast at occlusion boundaries. Photorealistic sources were designed to contain a combination of different challenges to evaluate performance on fine textures, complex occlusion areas of the scene, slanted planar surfaces, and continuous non-planar surfaces. One Multi-View (MV) algorithm and four light field algorithms (LF, Light Field Occlusion (LF-OCC), Epipolar Plane Image (EPI)2, EPI1) were used to evaluate the Mean Square Error (MSE), BadPix, and Bumpiness metrics and characteristics on each light field.

The "The 4D Light-Field dataset for material recognition" was proposed by Wang et al. [36] investigating how to differentiate and recognise materials on 4D light field. Differentiating materials in 2D images depends on their appearance in the image. The appearance of the objects depends on the shape and lighting. The authors introduced several novel Convolutional Neural Network (CNN) architectures specifically designed for material recognition on 4D light-field inputs. The best CNN architecture achieved about 6–7% boost compared with single 2D image material classification. The authors stated that this dataset was the first mid-size dataset for light-field images and it does appear that it is the first available material Light Field dataset. The dataset

[7] www.light field-analysis.net

(a) Input Image (b) Ground Truth

FIGURE 4. Full scene material classification examples [36].

contains 1200 images in 12 categories as follows: fabric, foliage, fur, glass, leather, metal, plastic, paper, sky, stone, water, and wood. The Lytro Illum camera was used to capture the images. The number of views of the images is 14×14 with resolution of 376×541. The main goal of this work was to investigate whether additional information in 4D light-field could improve the performance of material recognition over 2D images. In 2D images, the lighting reflectance of different objects is used to recognise the object type which is called surface reflectance or the (Bidirectional Reflectance Distribution Function (BRDF). In the light field image case, the alternative is to look at different views of a point. Convolutional Neural Network Visual Geometry Group 16 (CNN VGG 16) model was used in this test. Since this model was required to apply on 4D light field data, 5 different architectures (Viewpool, Stack, EPI, Angular Filter and 4D filter) were tested and their performance were compared. 2D average model was also performed for comparison purposes.

The fourth and final dataset reviewed in this section is the "Disney High Spatio-Angular Resolution Light Fields" [62] proposed by Kim et al. [33]. They proposed a method for scene reconstruction of indoor/outdoor complex static sources from 3D light fields. Their approach on scene reconstruction involved using a high but limited number of high resolution light field images captured with a DSLR camera mounted on a computer controlled 1.5 meter long Zaber T-LST1500D[8] linear stage. 100 images were acquired from each source. The spacing between each computer controlled camera position on the linear stage was in the range of 2mm to 15mm. The authors stated that by selecting this method of image acquisition, high spatio-angular resolution light fields were achieved whilst not taking the cost of full camera array or compromising either resolutions by single light field cameras based on lenslet array. The EPI derived from the light field images were used to generate precise depth maps of the static source and background as shown on Figure 5. The depth information was extracted for every visible scene point in the captured images. Although the authors concluded that their method could not compete with laser scanner in terms of absolute depth measurements, it resulted in higher resolution reconstruction.

[8]https://www.zaber.com/products/linear-stages/T-LST/details/T-LST1500D/features

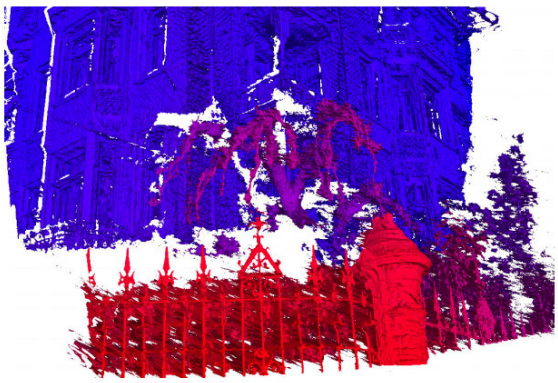


FIGURE 5. Shaded 3D mesh of an outdoor scene generated from triangulating individual depth maps [33].

Mousnier et al. [35] proposed an approach in reconstructing light field image using a dataset containing images from various indoor and outdoor sources. There are 30 images in total that were taken from 30 different sources of which 3 are with motion blur, one with long exposure time and one flat photograph. The authors calculated the depth map and the all-in-focus image from the focal stack, and then partially reconstructed the light field by the epipolar image using the depth map and the all-in-focus image.

C. QUALITY ASSESSMENT DATASETS

In this section, eight QoE datasets are reviewed. Tables 2, 3 and 4 illustrate respectively: the objective metrics that were considered in these datasets; the participants demography in the subjective tests; subjective assessment details relevant to each dataset.

Shi et al. [52] proposed the WIN-5LID dataset. The authors presented the first 5 DOF light field image dataset with the associated quality scores following visualisation on a stereoscopic display. This dataset primarily consists of 6 real source images from [14] and 4 synthetic sources from [37]. The authors introduced different levels of distortions on the primary LF images to experiment different levels of visual experiences. The authors used in particular 22 Hypothetic Reference Circuits (HRC) from two compression methods (High Efficiency Video Coding (HEVC) with 5 different

Quantisation Parameter (QP)s and JPEG2000 with 5 different Compression Ratio (CR)s) and 4 reconstruction algorithms in the angular domain (Linear Interpolation (LI) and Nearest Neighbour interpolation (NNI) with 5 different sub-sampling rates and two state-of-the-art CNN models). As a result, a dataset of 220 processed images was created.

In subjective assessment a light field image quality was compared against its sub-views. Since the quality of a light field image is influenced by some inherent properties like motion parallax and refocusing, the expectation was to observe a difference between a sub-view (picture) quality and the quality of light field (overall). The subjective assessment results show that the picture quality proves to be better than overall quality across all HRCs. Subjective assessment details are listed in Table 2.

The 19 objective metrics that were tested in objective assessment are listed in Table 4. Four correlation measures were used to evaluate the correlation between MOS and predicted results. Spearman Rank Order Correlation Coefficient (SROCC) and Kendall's rank-order correlation coefficient (KROCC) were used to measure the monotonicity of MOS prediction. Pearson Linear Correlation Coefficient (PLCC) and Root mean-squared error (RMSE) methods were used to measure the linear relationship between the two sets and the prediction accuracy, respectively. The results show that S3D integrated quality (SINQ) performs better in picture quality with high positive correlations of 0.8890, 0.9314 and 0.7364 in SROCC, PLCC and KROCC and lowest value of 0.3922 in RMSE. In overall quality, however, Information content weighted structural similarity measure (IW-SSIM) provides the best results with high positive correlations of 0.8352, 0.8435 and 0.6542 in SROCC, PLCC and KROCC and lowest value of 0.5492 in RMSE. The results for objective and subjective metrics are not well correlated as objective metrics do not consider the intrinsic characteristics of light field images.

Viola and Ebrahimi [51] proposed the VALID dataset. Five LF plenoptic images from [14] were selected for this dataset. Using the Matlab Toolbox v0.4, the authors generated 15×15 perspective views of 625×434 pixels and depth of 10 bits per color channel for each image. The authors considered two output bit depths - 10 bits and 8 bits - for their objective and subjective tests. The reason they chose 8 bits per colour channel was to allow their test to be compatible with the current consumer market displays. Among the distortion methods listed in Table 2, only HEVC and VP9 were applied on 8 bit colour channel outputs.

For objective quality assessment, the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) objective metrics were considered to assess the visual quality of the content. The objective metrics were calculated on both 8 bit and 10 bit depth outputs. However, as mentioned in Table 4, the authors did not specify which metric resulted better in their objective assessment.

Two different subjective visualisation assessment methodologies were considered for 8 bit and 10 bit outputs. For 10 bit output contents, passive methodology without refocussing was performed. The stimulus was presented to participants alongside the reference image in a side-by-side arrangement. Participants would compare and give their mark from -3 (much worse) to +3 (much better) with 0 indicating no preference. In case of 8 bit output, three methodologies were considered: passive, interactive and the combination of both. For the"passive and interactive" method, the participants were presented an animation and were only allowed to give their score at the end of the 13.6 s animation. A five-point scale was used for participants to feedback their visual experience. In this method, refocusing was also adopted to enable participants to refocus foreground and background at a rate of 4 frames per second. The authors did not provide information about the demography of participants who partook in their subjective test.

Tamboli et al. [49] proposed a high angular resolution turntable quality dataset using three real life objects as shown in Figure 6, a turntable and a single camera (see Table 1 for details). The angular resolution was set to one view per degree. This meant the authors captured one image per degree, hence 360 images were captured from each object. The camera was placed in such a way that its principal axis was perpendicular to the rotation axis of the turntable. The authors stated that the reasons they decided to create their own dataset using a single camera and a turntable to perform the quality test were as follows: (1) the existing turntable datasets at the time had angular resolution of more than 1 degree; (2) where a dataset did meet the above resolution, there were more than one object on the turntable; (3) the available datasets used artistic objects that exhibited non-Lambertian surface; (4) where the dataset had a single object on the turntable, the centre of the object was not aligned with the vertical axis of the rotation. Therefore, in capturing images using turntable, the centre of the object was aligned with the axis rotation of turntable.

The Holovizio HV721RC display generated 3D views from subsets containing turntable images. In subjective tests, subjects were allowed to change position along an arc marked in front of the 3D display; however, the evaluation of the 3D content were performed at 5 specifically selected positions along the arc. For subjective tests, 5 different levels of distortions were applied to 8 different views from each object, resulting in a per object total of 40 distorted 3D views. In addition to distorted images, 8 reference views were also shown for each object which were presented randomly to subject twice. Therefore in total 168 images were selected and shown to participants in the subjective test. The conversion of 3D view images to light field image was done using proprietary light field convertor provided with the display. Subjective evaluation was carried out from 5 different positions as shown in Figure 7 within the display's Field of View (FoV) in an 8ft radius distance from the display. Each location was 12.5 degrees apart from the adjacent position. As described in Table 3, 20 participants assessed the content on a five-point scale, with 1 corresponding to poor quality

TABLE 2. QoE datasets - subjective assessment details.

Dataset	Representative Artifacts	Distortions Methods	Reconstruction Methods	Subjective evaluation method	Scores provided [All subject scores + MOS / MOS only]	Display Type	Display Make/Model
SMART [48]	Compression	JPEG, JPEG2000 HEVC Intra SSDC	NA	Pairwise Comparison Matrix (PCM), using Bradley-Terry (BT) model to convert discrete rating data in matrix to continuous rating scale	All subject scores + MOS	2D	Dell U2413f
VALID [51]	Compression Refocusing	HEVC, VP9 3 Other Methods	NA	DSIS, 5 points scale for 8-bit Output Depth CB-ACJ	All subject scores + MOS	2D	Eizo ColorEdge CG318-4K Samsung SyncMaster2443
Turntable Quality Dataset [49]	Compression, blur, additive noise	Additive Gaussian Noise JPEG compression Gaussian blur	NA	SSCQE, 5 points scale at 5 different locations within display's FOV	All subject scores + MOS	3D Light field	Holografika's Holovizio HV721RC
LFDD: Annotated LF Dataset [54]	Compression Noise Geometric Distortion Contrast Enhancement	JPEG, JPEG2000, BPG VP9, AV1, AVC, HEVC, noise, Geometric Distortions	NA	DSIS, points between 1 to 5 with one decimal precision	MOS only	NA	Crowdsourced
WIN-5LD [52]	Compression Reconstruction Refocussing	JPEG2000 HEVC	Linear Interpolation Nearest Neighbour 2 CNN based methods	DSCQS, 5 points scale	MOS only	3D Stereoscopic	55" Samsung 3D TV
MPI-Light Field [50]	Compression Reconstruction	3D-HEVC Nearest Neighbour [NN] Image wraping OPT Quantised Depth Maps (DQ) Display: Gaussian Blur in angular domain	Linear Nearest Neighbour [NN] Optical Flow Estimation [OPT]	JND and JOD, scale between 0 to -9 based on severity of distortion	All subject scores + MOS	2D (used with active shutter glasses)	ASUS VG278 27" FULL HD LCD desktop monitor together with NVIDIA 3D glasses
Shan et al. [53]	Compression, blur, additive noise	Gaussian Blur JPEG JPEG2000 White Noise	NA	DSCQS, 5 points scale	All subject scores + MOS	2D	Dell E2211Hb
V-SENSE [55]	NA	NA	Refocusing	Eye Tracking, recording events, saccades, fixations and blinks of left eye only	NA	2D	Dell P2415Q 23.8" Monitor

TABLE 3. QoE datasets - participants demography.

Dataset	Participants	Male	Female	Comments
SMART [48]	19	NA	NA	UG/PG Students - Naïve about image rendering/processing
VALID [51]	NA	NA	NA	NA
Turntable Quality Dataset [49]	20	NA	NA	Students unfamiliar with the process of image rendering
LFDD: Annotated LF Display [54]	16	8	8	Naïve participants aged between 15 and 57 years
WIN-5LD [52]	23	14	9	29 non-expert participants aged between 19 and 26
MPI-Light Field [50]	40	20	20	Aged between 24 and 40 years with normal or corrected-to-normal vision.
Shan et al. [53]	19	14	5	Aged between 22 and 35 years
V-SENSE [55]	21	16	5	Aged between 18 and 37 years

TABLE 4. Distribution of objective assessment metrics used in each QoE datasets.

Dataset	2D-SSIM	3OMS-SSIM	3D-SSIM	3DSWIM	BRISQUE	BSVQE	Chen	FSIM	FSIMc	GMSD	HDR-VDP-2	IFC	IW-PSNR	IW-SSIM	MP-PSNR	MSE	MS-SSIM	NFERM	NIQE	NQM	PSNR	PSNR-HVS	PSNR-HVS-M	SINQ	SIQM	SNR	SSIM	STSDLC	UIQI	UQI	VIF	VIFP	VSNR	WSNR	Q_{3D}	Best Correlation
SMART [48]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	✓	✓	✓	-	-	✓	✓	-	-	✓	-	✓	✓	✓	-	PSNR-HVS-M
VALID [51]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	Not Specified
Turntable Quality Dataset [49]	-	-	-	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	✓	-	-	-	✓	-	-	-	-	-	✓	-	-	-	-	-	-	-	✓	Q_{3D}
LFDD: Annotated LF Display [54]	-	✓	-	-	-	-	-	✓	✓	-	-	-		-	-	✓	-	-	-	-	✓	-	-	-	-	-	✓	-	-	✓	✓	-	-	-	-	FSIM, FSIMc & VIF
WIN-5LD [52]	-	-	-	-	✓	✓	✓	✓	-	-	-	✓	✓	✓	-	-	✓	✓	✓	✓	✓	-	-	✓	-	-	✓	-	-	✓	✓	✓	✓	✓	-	SINQ in picture quality IW-SSIM in overall quality
MPI-Light Field [50]	✓	-	✓	✓	-	-	-	-	-	✓	✓	-	-	✓	✓	-	-	-	-	-	✓	-	-	-	✓	-	-	✓	-	-	-	-	-	-	-	HDR-VDP-2 & GMSD
Shan et al. [53]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Verified
V-SENSE [55]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NA

and 5 corresponding to excellent quality. The age range and gender distribution of participants were not mentioned by authors.

The authors used six objective assessment metrics, as illustrated in Table 4, of which Multiscale Structural Similarity (MS-SSIM) achieved the highest values for Linear Correlation Coefficient (LCC) and SROCC.

They also proposed Q_{3D}, a new 3D full reference objective quality metric that considers the spatio-angular nature of 3D content and aligns very closely with subjective test scores. This metric was also tested on unseen data (a single synthetic image affected by the same distortion types and levels). Subjective tests with 20 participants were also carried out on this new set of images and Difference Mean Opinion Scores (DMOS) were calculated from subjects' opinion scores. The authors' proposed 3D objective quality metric proved to evaluate 3D views, rather than each contributing 2D component of 3D view; tested on the single new synthetic content, it achieved a value of 0.877 for LCC whereas the highest value achieved by 2D quality metrics was 0.759.

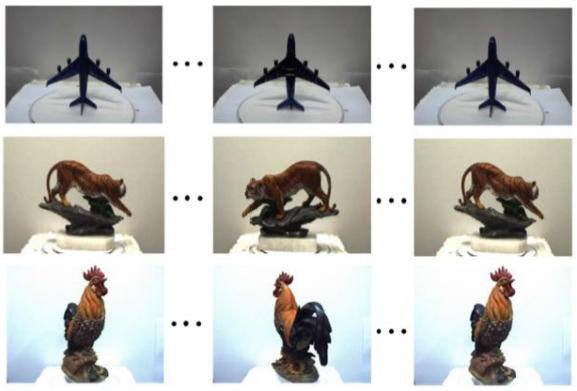


FIGURE 6. Three different views of the three real objects used in generating the turntable dataset [49].

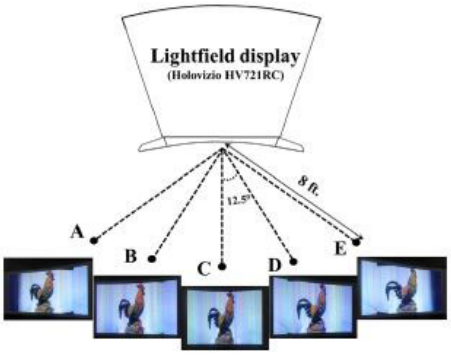


FIGURE 7. Turntable dataset [49] - Five distinct positions within display FoV for participants in Subjective Test.

Zizen and Fliegel [54] proposed the purely synthetic Annotated Light Field Image Dataset (LFDD). They used 10 LF images of which 8 synthetic images were obtained from [37], investigated in the previous section. The other two images were used in training sessions of subjective evaluation. As listed in Table 2, five degradation levels were chosen for each distortion method. The authors mainly concentrated on applying different types of compression methods as means of testing distortions on Light Field images. They stated that, due to the fact that compression will have to play a major part in huge size light field images, it is necessary to pay special attention on compression algorithms and the artifacts that are exhibited by the algorithms. However, the authors acknowledged the fact that compression artefacts are not the only type of distortion that could have an impact on the perceived quality in subjective tests. Geometrical distortions and distortions during the acquisition process would not show up during the encoding process but could also have an impact on subjective quality test. The PVSs (via different compression methods, noise, etc.) were obtained via Matlab[9] or FFMpeg.[10]

Subjective tests were conducted through crowdsourcing. A modified version of the DSIS methodology was deployed for this subjective test, with a different score scale. Instead of a 5-point integer scale, the observers were able to give a score in the precision of one decimal place. The serpentine scan order was chosen by the authors to accommodate the sources into a pseudo-sequence and played the sequence as animation to observers. The reference source was also presented to observers simultaneously to the stimulus in a side by side set up. The degradation in video based codecs were more noticeable than image based codecs due to the selected lower bit rates for different levels of video based codecs.

Significant changes were observed by comparing MOS of subjective tests with the results of objective metrics as listed in Table 4, i.e., impulse noise distortion gave better scores in objective scores than subjective tests, while geometrical distortions evaluated better in subjective tests. For image-based compression algorithms, the Feature Similarity Index for Image (FSIM) and FSIM colour (FSIMc) metrics performed better than the VIF metric. However the VIF metric showed a strong correlation with subjective results for video encoders and all the compression algorithms combined.

Shan et al. [53] proposed a dataset which consists primarily of six light field images from the EPFL Light-Field Image Dataset [13] that was investigated earlier in this paper. For each LF image they created an animation of all viewpoints - all focus view- of each LF image where each view was a frame. The frame rate of the animation was set to 15 frames per second. They proposed a new subjective evaluation method for LF images and compared their experimental result with PSNR results. Two LF images were presented side by side in the subjective test on the 2D display listed in Table 2. Each original light field image underwent 4 different distortion methods, listed in Table 2, and for each distortion method, 4 different levels were considered. Therefore for each SRC two sets of 16 frames animations were generated for subjects to observe and rate. The spatial resolution of each frame of the animation is 625×434. The Double Stimulus Continuous Quality Scale (DSCQS) subjective rating methodology was adopted and subjects were given the chance to rate the quality of the images they were observing on a scale of 1 to 5 with 1 being ''Bad'' quality and 5 being ''Excellent'' quality. The authors noted that, from subjective test results and studying the MOS against distortion level for test contents, the distortion level of compression distortion (JPEG and JPEG2000) was difficult to distinguish, but the contents of white noise and Gaussian blur were easier for evaluation. The authors considered adding more distortion in their future work to complete the LF quality assessment database. Furthermore the subjective test results showed that the results depended on the texture and colorfulness of the test images. Objective test results were close to the subjective test results. The authors concluded that their method of using all of the all-in-focus views of LF image to produce animation sequences as test data helped to fully reflect the quality of LF images.

Adhikarla et al. [50] proposed a dense light field dataset, called the MPI-Light Field dataset, that consists of 9 synthetic

[9]https://www.mathworks.com/
[10]https://ffmpeg.org

sources and 5 real world sources. The real world sources were captured considering horizontal parallax only, using a digital 2D camera, as listed in Table 1, mounted on a one-meter long motorized linear stage. The spatial and angular resolution of all light fields in this dataset are $960 \times 720 \times 101$, i.e., 101 2D views per source, hence a total of 1,414 2D reference images across 14 sources. Four reconstruction methods (with six angular subsampling factors), six compression levels (with 3D extension of HEVC encoder) and six display distortion (modeled using Gaussian blur in the angular domain) levels, as illustrated in Table 2, were applied to the reference light fields. NNI, LI, Optical Flow Estimation (OPT), and Quantized Depth Maps (DQ) algorithms were applied on synthetic sources while NNI, OPT, Gaussian blur (GAUSS), and HEVC were applied on real world sources. As a consequence of the application of the distortion methods listed in Table 2 at different severity levels, the dataset contains 350 different light fields. Figure 8 reports an example for each type of distortion considered.

For subjective assessment the authors used a 2D desktop monitor with a 3D vision kit (with active shutter glasses) for displaying stereoscopic images. Motion parallax was reproduced using a custom head tracking. The authors adopted a new approach in scaling pair-wise comparison results, using Just Objectionable Differences (JOD), since the interest was in identifying different qualities rather than visible differences that may result in similar qualities.

The authors also conducted objective assessment, testing existing metrics on the dataset. The metrics used in their assessments, listed in Table 4, were applied on individual light field images and then the scores over all images were averaged. The results of objective quality metrics showed good performance of 2D and video quality metrics. Some metrics did not perform well on some specific distortions; for example HDR-VDP-2, although performing very well in general, did not perform well in predicting the quality for HEVC distorted images. The authors concluded that training such metrics for a particular distortion type could boost their performance. Another result was that proposed extensions of the SSIM metric, aiming at accounting for angular changes, did not perform well. This area requires more investigation and development.

Paudyal et al. [65] and [48] proposed the SMART dataset. 16 sources were selected for this dataset considering some key features in image quality such as Spatial Information (SI), Colorfulness (CF), contrast, correlation, homogeneity, brightness, hue, and saturation. The dataset also includes the depth map images for each scene. Figure 9 presents three sources from this dataset. The raw light fields captured by the Lytro Illum camera were then processed via Matlab Toolbox v0.4 to generate 2D light field images with 2D spatial resolution of 434×625, and 15×15 views and RGB and an additional weighting channel. The authors evaluated the perceptual quality of compressed light fields through subjective quality assessment. Four different encoding methods as listed in Table 2 at different levels were investigated in this subjective assessment.

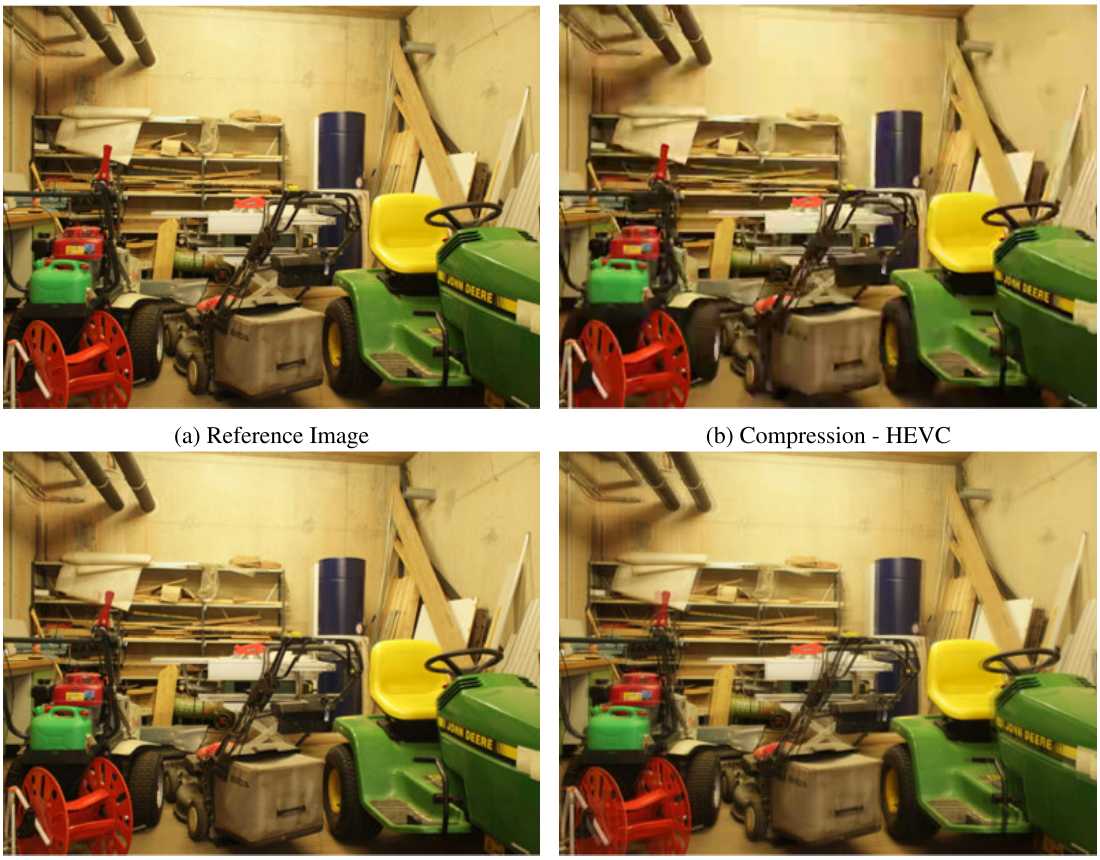
Subjective tests were conducted with 19 Undergraduate (UG)/Postgraduate (PG) students who were naïve about image rendering/processing. The age range and gender distribution of the participants were not mentioned by the authors. The pair-wise evaluation method was adopted for subjective assessment in which participants were presented a pair of images on a designed graphical user interface, on a 2D display as listed in Table 2, to choose the best quality one as per their perception Just Noticeable Differences (JND). The subjective assessment results highlighted the difference in perceptual quality of different encoding methods, with the plenoptic image compression method Sparse Set and Disparity Coding (SSDC) achieving the highest score amongst the four methods.

Objective quality assessment was performed considering the objective metrics listed in Table 4. The correlation between the subjective opinion scores and objective quality scores was evaluated using PLCC, SROCC, and Kendall Tau rank Correlation Coefficient (KTCC). Amongst all objective metrics, PSNR-HVS Visual Masking (PSNR-HVS-M) achieved higher scores in terms of PLCC and SROCC which could be associated with the fact that PSNR-HVS-M takes into account the Human Visual System (HVS) and the contrast sensitivity of the image.

The authors observed that the level of compression (for a fixed QP) depended on the image content and in particular on color levels and brightness. However, no significant correlation was identified between the considered descriptor and the subjective quality of the light field images.

The last QoE dataset we cover in this section is V-SENSE, proposed by Gill et al. [55]. The dataset was developed in the framework of a QoE study to investigate the visual perception of light field from the users' perspective. The investigation focused on whether the user attention on light field content was different from 2D content and the impact of varied focus on viewer attention. The authors used the refocusing method on light field content to examine how it impacted the visual attention of participants.

This dataset consists of twenty light fields selected from four publicly available datasets [12], [13], [33], and [37], all of which have been reviewed in this paper. The light field selection criteria from the four datasets were light fields with multiple regions or objects with high colour contrast and high edge density and local contrast at varied depths and spatial locations. The Fourier disparity layers method [66] was used to generate focal stacks for light fields in various focus scenarios such as: (1) all-in-focus, where all points in the rendered image were in focus equally; (2) region-in-focus, where only one image of the focal stack was rendered, hence objects in the light field with the depth of focus of that particular slice would appear sharp; (3) focal-sweep, where all images of the focal stack were rendered in a sequence.



(a) Reference Image (b) Compression - HEVC

(c) Reconstruction Interpolation (d) Gaussian Blur

FIGURE 8. The impact of different distortion methods on image quality - MPI-Light Field [50].

To study the participants' area of attention, participants' eye tracking data were collected using the Eyelink 1000 plus [67] desktop mounted eye-tracker. The eye tracker only recorded the movement of the participants' left eye. The subjective test was conducted to study how the participants viewed the frames they were presented and which region of either all-in-focus, region-in-focus (left region or right region of the image) or the focal sweep (front-to-back or back-to-front) drew more attention. Saliency maps from the fixation points recorded by the eye-tracker were used. The results showed that, whilst in some cases the salient points were the same in different rendering patterns, objects could draw participants' attention in focal sweep rendering but not in all-in-focus scenario. Furthermore the authors found that if an object is salient in all-in-focus and becomes in focus in any of the other two rendering scenarios, it becomes salient too. When an object had a high level of saliency in all-in-focus rendering and was observed in another rendering method where it was not in the focus region, even though the object was not in focus it still pulled viewers' attention from the focus area. The saliency patterns in different rendering methods were very much dependent on the nature of the objects in the image. Images with a high number of distinct objects like a treasure chest full of jewels demonstrated a clear dispersed saliency pattern where participants tended to follow the region in focus.

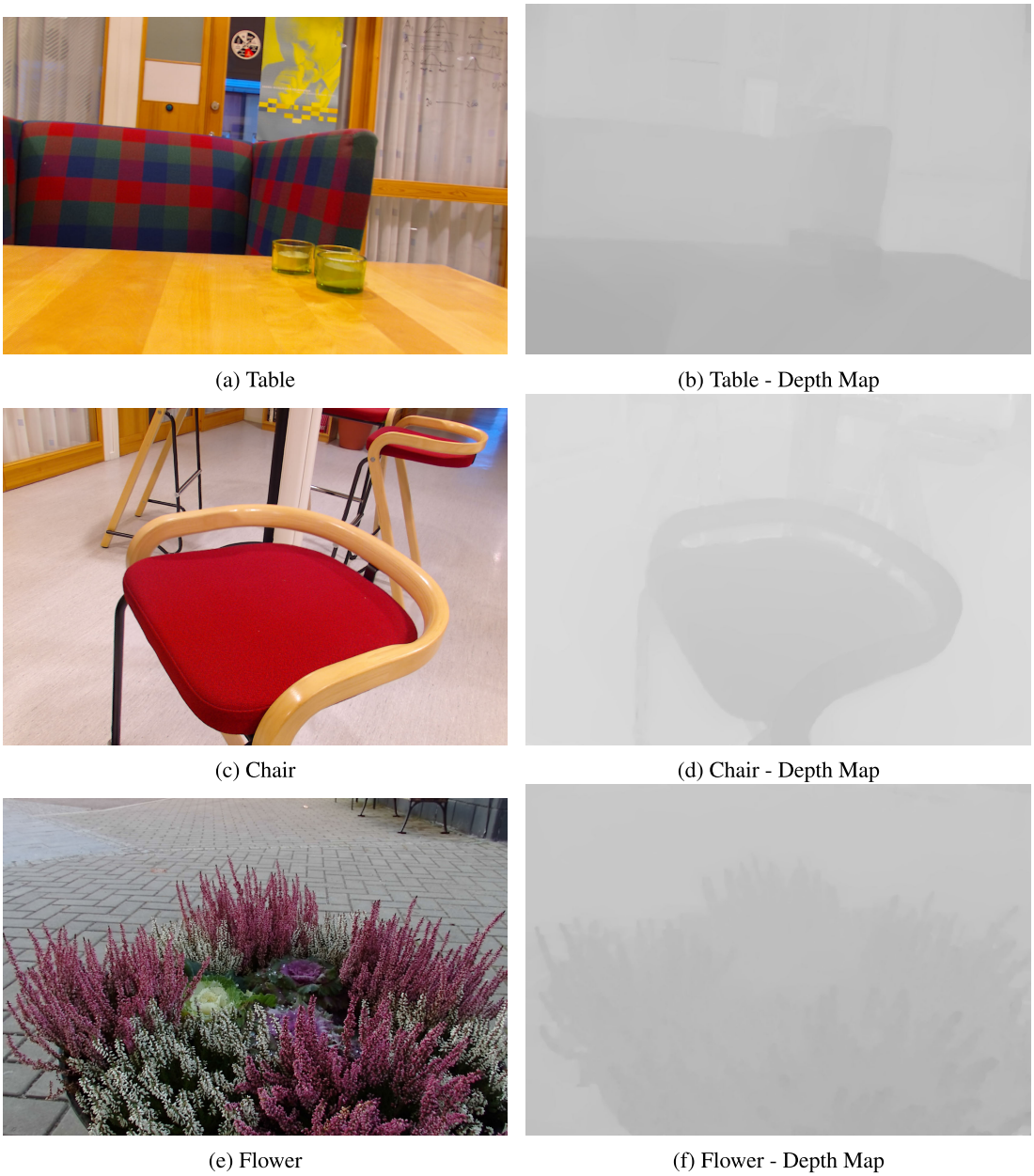
III. CHALLENGES AND FUTURE DIRECTIONS

A. DATASETS SOURCE CONTENT

The datasets covered in this paper can be grouped into synthetic light fields, real world light fields captured with a lenslet light field camera (Lytro Illum or Raytrix) and real world sources captured with a camera array or gantry (or a mix of both) or a fixed single camera with content on a turntable.

The challenge in using plenoptic cameras such as Lytro Illum is that, when using software tools such as the MATLAB Light Field Toolbox for decoding the recorded light field information, part of the color and depth information could be distorted. Even though the Toolbox's color correction function could be used to correct the impact, the generated distortion could still have a negative impact on the perceived quality.

Furthermore, based on the aim of the research, the type and variety of the light fields in the dataset would need to be considered. Datasets containing real-world sources or with a large proportion of real world sources are often required. For instance, for evaluating compression algorithms, such



(a) Table (b) Table - Depth Map

(c) Chair (d) Chair - Depth Map

(e) Flower (f) Flower - Depth Map

FIGURE 9. Three scenes from the SMART Dataset [48] with corresponding Depth Maps.

as in [48], real world source datasets would be required. Indeed this would also hold true when working on the design and performance evaluation of objective quality metrics. This would necessitate real world source datasets with a variety of light field contents, contrast and other features. Synthetic datasets are required for instance when the need of highly accurate and predefined ground truth depth or disparity maps comes into force. In addition, very few datasets provide video data, with most of them focusing on static sources. Acquiring LF video data requires plenoptic cameras or arrays of cameras, expensive and requiring appropriate lab environments in the second case, while static sources can be recorded via moving a single camera in multiple positions or moving the object in front of a single camera, e.g. via a turntable. More datasets with video data should be produced and made public to advance video-related research and testing. In general, the scientific community would need more datasets with different contents and also with associated subjective scores associated to visualisation in light field displays.

B. OBJECTIVE METRICS

As reviewed above in the part on objective quality assessment of the light field data in the datasets, existing objective quality metrics do not result in high correlation with subjective quality assessment results, in particular when tested on new datasets. This is on one hand due to the fact that

the traditional 2D objective quality metrics do not consider specific features of light field images such as motion parallax and refocusing. This requires more research in this area of designing specific objective quality metrics for light field images.

Another point that can be made observing Table 4 is that there is no single best objective quality metric that always correlates best with subjective assessment results. Based on the nature of the light field and its journey from camera to display, the correlation of the objective quality metric with the subjective results changes. Again, the reason behind this variety could be the lack of light field specific objective quality metrics in the datasets and the fact that most of them have been designed based on a limited set of data.

It is worth mentioning that quality metrics developed specifically for light fields exist (see e.g. [68], [69], [70], [71], [72]). Some of them, however, were developed and tested based on quality scores of subjects visualising content in commercial displays not designed for the visualisation of light fields. A comparison on existing and new objective metrics based on datasets with light field display visualisation quality scores would indeed be important.

C. SUBJECTIVE ASSESSMENT

One of the current problems in subjective assessment is the lack of available light field displays in the research community. From the eight QoE dataset works that we have covered in this paper, only one used a 3D light field display. To achieve the best end user perceptual quality and to measure the actual light field experience, an appropriate display needs to be used. Subjective tests using light field displays have indeed been performed in the past and their results are reported in the literature (see e.g. [59], [60], [61], [73], [74]), although the relevant datasets are not public. A summary of the subjective tests performed with light field displays is provided in Table 1 in [75].

In additions, agreed guidelines for subjective assessment for light field visualisation are missing. Initial considerations on specific aspects, i.e. viewing distance, are reported in [76]. The IEEE Standard for the Quality Assessment of Light Field Imaging (P3333.1.4) [77] is addressing these aspects.

Finally, most of the existing quality datasets report the aggregated scores from subjects, via MOS. However, the value of including scores from all subjects and not only MOS has recently been established [78], [79], [80], [81], [82], as the potential bias of each subject in assessing the content and the interaction with content can be considered. Indeed, the aim of subjective tests is to estimate the ''population mean opinion score'' having availability of a limited pool of subjects in the population for performing the tests. Reporting all the subjects' scores on the different contents will enable a more accurate estimation of the ''population mean opinion score'', for instance via the model in [78].

For further discussion on subjective assessment for light field displays, as well as a summary of light field subjective tests using light field displays, readers can refer to [75].

IV. CONCLUSION

This paper presented the review of 33 publicly available datasets, eight of which for the purpose of quality assessment. For the latter, the paper compares the objective and subjective assessment methods used. The challenges in creating datasets and the fact that different types of datasets are required for different researches are discussed. The need for a robust light field objective quality metric is also discussed.

We expect this work will support research scientists working in the areas of light field compression, transmission, post-processing, rendering and quality evaluation in the selection of the most appropriate datasets for developing and testing their algorithms. The taxonomy and collection of the available data provided here will also support the development of machine learning strategies for the quality assessment of light field imaging, requiring large amount of data representing different contents and use cases.

A major issue highlighted is the lack of datasets where subjective scores are provided following visualisation on light field displays: most datasets use 2D or stereoscopic 3D displays, with only one dataset with subjective tests on a light field display (horizontal parallax). We expect that the lowering costs of light field displays will support the development of new datasets. We plan to provide relevant datasets and we also solicit the research community to make datasets publicly available. This will underpin the development of novel quality assessment metrics and compression methods. For video quality assessment, there is a lack of relevant light field datasets (video content) that is preventing the development and testing of video quality metrics for light field content. Also, with the development of full parallax displays, datasets with suitable content and with subjective tests on full parallax displays would be extremely beneficial for the research community in the area. More types of distortions specific for light field content should also be considered, as most of the QoE datasets only consider compression distortion in addition, in some cases, to Gaussian noise and blur.

REFERENCES

[1] M. Levoy and P. Hanrahan, ''Light field rendering,'' in *Proc. 23rd Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 1996, pp. 31–42.

[2] A. Gershun, ''The light field,'' *J. Math. Phys.*, vol. 18, nos. 1–4, pp. 51–151, 1939.

[3] E. H. Adelson and J. R. Bergen, *The Plenoptic Function and the Elements of Early Vision*, vol. 2. Cambridge, MA, USA: Vision and Modeling Group, Media Laboratory, Massachusetts Institute f Technology, 1991.

[4] G. Wu, ''Light field image processing: An overview,'' *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 926–954, Oct. 2017.

[5] P. A. Kara, A. Cserkaszky, M. G. Martini, A. Barsi, L. Bokor, and T. Balogh, ''Evaluation of the concept of dynamic adaptive streaming of light field video,'' *IEEE Trans. Broadcast.*, vol. 64, no. 2, pp. 407–421, Jun. 2018.

[6] L. Shi, W. Zhou, Z. Chen, and J. Zhang, ''No-reference light field image quality assessment based on spatial-angular measurement,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 4114–4128, Nov. 2020.

[7] Y. Tian, H. Zeng, J. Hou, J. Chen, J. Zhu, and K.-K. Ma, "A light field image quality assessment model based on symmetry and depth features," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 5, pp. 2046–2050, May 2021.
[8] W. Wen, K. Wei, Y. Fang, and Y. Zhang, "Visual quality assessment for perceptually encrypted light field images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 7, pp. 2522–2534, Jul. 2021.
[9] Q. Zhang, S. Wang, X. Wang, Z. Sun, S. Kwong, and J. Jiang, "A multi-task collaborative network for light field salient object detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 5, pp. 1849–1861, May 2021.
[10] Y. Zhang, W. Dai, M. Xu, J. Zou, X. Zhang, and H. Xiong, "Depth estimation from light field using graph-based structure-aware analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 4269–4283, Nov. 2020.
[11] W. Ellahi, T. Vigier, and P. L. Callet, "Analysis of public light field datasets for visual quality assessment and new challenges," in *Proc. Eur. Light Field Imag. Workshop*, 2019, pp. 2–6.
[12] V. Vaish and A. Adams, "The (new) Stanford light field archive," Comput. Graph. Lab., Stanford Univ., Stanford, CA, USA, Tech. Rep. 7, 2008, vol. 6, no. 7.
[13] M. Rerabek, L. Yuan, L. A. Authier, and T. Ebrahimi, *ISO/IEC JTC 1/SC 29/WG1 Contribution EPFL Light-Field Image Dataset*, document ISO/IEC JTC 1/SC 29/WG1, 2015.
[14] M. Rerabek and T. Ebrahimi, "New light field image dataset," in *Proc. 8th Int. Conf. Quality Multimedia Exper. (QoMEX)*, 2016, pp. 1–2.
[15] *Synthetic Light Field Archive*. Accessed: Sep. 27, 2021. [Online]. Available: https://web.media.mit.edu/~gordonw/SyntheticLightFields/index.php
[16] *Gordon Wetzstein*. Accessed: Sep. 27, 2021. [Online]. Available: http://web.stanford.edu/~gordonwz/index.html
[17] *Static Planar Light-Field Test Dataset*. Accessed: Sep. 27, 2021. [Online]. Available: https://www.iis.fraunhofer.de/en/ff/amm/dl/lightfielddataset.html
[18] L. Guillo, X. Jiang, G. Lafruit, and C. Guillemot, *Light Field Video Dataset Captured by a R8 Raytrix Camera*, document ISO/IEC JTC1/SC29/WG1 & WG11, 2018.
[19] R. R. Tamboli, M. S. Reddy, P. A. Kara, M. G. Martini, S. S. Channappayya, and S. Jana, "A High-angular-resolution turntable data-set for experiments on light field visualization quality," in *Proc. 10th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2018, pp. 1–3.
[20] *Comprehensive Light Field Image Dataset*. Accessed: Jul. 10, 2022. [Online]. Available: https://http://lightfield.stanford.edu/
[21] S. Shekhar, S. Beigpour, M. Ziegler, M. Chwesiuk, D. Palen, K. Myszkowski, J. Keinert, R. Mantiuk, and P. Didyk, "Light-field intrinsic dataset," in *Proc. Brit. Mach. Vis. Conf.*, 2018, p. 120.
[22] *The Stanford Multiview Light Field Datasets*. Accessed: Sep. 27, 2021. [Online]. Available: http://lightfields.stanford.edu/mvlf/
[23] D. G. Dansereau, B. Girod, and G. Wetzstein, "LiFF: Light field features in scale and depth," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8042–8051.
[24] F. S. Zakeri, A. Durmush, M. Ziegler, M. Batz, and J. Keinert, "Non-planar inside-out dense light-field dataset and reconstruction pipeline," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1059–1063.
[25] *TEDDY: A High-Resolution High Dynamic Range Light-Field Dataset*. Accessed: Sep. 27, 2021. [Online]. Available: https://fordatis.fraunhofer.de/handle/fordatis/158
[26] M. S. Khan Gul, T. Wolf, M. Batz, M. Ziegler, and J. Keinert, "A high-resolution high dynamic range light-field dataset with an application to view synthesis and tone-mapping," in *Proc. IEEE Int. Conf. Multimedia Expo. Workshops (ICMEW)*, Jul. 2020, pp. 1–6.
[27] *Densely Sampled Light Field Datasets*. Accessed: Sep. 27, 2021. [Online]. Available: https://civit.fi/densely-sampled-light-field-datasets/
[28] S. Moreschini, F. Gama, R. Bregovic, and A. Gotchev, "CIVIT dataset: Horizontal-parallax-only densely-sampled light-fields," in *Proc. Eur. Light Field Imag. Workshop*, vol. 6, 2019, pp. 1–4.
[29] *A Benchmark of Light Field View Interpolation*. Accessed: Sep. 27, 2021. [Online]. Available: https://www.cl.cam.ac.uk/research/rainbow/projects/lightfield-benchmark/
[30] D. Yue, M. S. Khan Gul, M. Batz, J. Keinert, and R. Mantiuk, "A benchmark of light field view interpolation methods," in *Proc. IEEE Int. Conf. Multimedia Expo. Workshops (ICMEW)*, Jul. 2020, pp. 1–6.
[31] *4DLFVD: A 4D Light Field Video Dataset*. Accessed: Sep. 27, 2021. [Online]. Available: https://ieee-dataport.org/open-access/4dlfvd-4d-light-field-video-dataset
[32] X. Hu, C. Wang, Y. Pan, Y. Liu, Y. Wang, Y. Liu, L. Zhang, and S. Shirmohammadi, "4DLFVD: A 4D light field video dataset," in *Proc. 12th ACM Multimedia Syst. Conf.*, Jun. 2021, pp. 287–292.
[33] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. H. Gross, "Scene reconstruction from high spatio-angular resolution light fields," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1–73, 2013.
[34] S. Wanner, S. Meister, and B. Goldluecke, "Datasets and benchmarks for densely sampled 4D light fields," in *Proc. VMV*, vol. 13, 2013, pp. 225–226.
[35] A. Mousnier, E. Vural, and C. Guillemot, "Partial light field tomographic reconstruction from a fixed-camera focal stack," 2015, *arXiv:1503.01903*.
[36] T.-C. Wang, J.-Y. Zhu, E. Hiroaki, M. Chandraker, A. A. Efros, and R. Ramamoorthi, "A 4D light-field dataset and CNN architectures for material recognition," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 121–138.
[37] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, "A dataset and evaluation methodology for depth estimation on 4D light fields," in *Proc. Asian Conf. Comput. Vis.*, Cham, Switzerland: Springer, 2016, pp. 19–34.
[38] *Learning to Synthesize a 4D RGBD Light Field From a Single Image*. Accessed: Sep. 27, 2021. [Online]. Available: https://github.com/pratulsrinivasan/Local_Light_Field_Synthesis
[39] P. P. Srinivasan, T. Wang, A. Sreelal, R. Ramamoorthi, and R. Ng, "Learning to synthesize a 4D RGBD light field from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2243–2251.
[40] *Synchronized Light-Field Video*. Accessed: May 23, 2022. [Online]. Available: https://www.interdigital.com/data_sets/light-field-dataset
[41] N. Sabater, G. Boisson, B. Vandame, P. Kerbiriou, F. Babon, M. Hog, R. Gendrot, T. Langlois, O. Bureller, A. Schubert, and V. Allie, "Dataset and pipeline for multi-view light-field video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 30–40.
[42] *Matching Light Field Datasets From Plenoptic Cameras 1.0 and 2.0*. Accessed: Sep. 27, 2021. [Online]. Available: https://figshare.com/articles/dataset/The_Plenoptic_Dataset/6115487/3
[43] W. Ahmad, L. Palmieri, R. Koch, and M. Sjostrom, "Matching light field datasets from plenoptic cameras 1.0 and 2.0," in *Proc. 3DTV-Conf., True Vis. Capture, Transmiss. Display 3D Video (3DTV-CON)*, Jun. 2018, pp. 1–4.
[44] L. Palmieri, R. Koch, and R. O. H. Veld, "The plenoptic 2.0 toolbox: Benchmarking of depth estimation methods for MLA-based focused plenoptic cameras," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 649–653.
[45] *The Plenoptic 2.0 Toolbox*. Accessed: Jan. 77, 2022. [Online]. Available: https://zenodo.org/record/3558284#.YeXpMHrP2Hs
[46] M. Schambach and M. A. Heizmann, "Multispectral light field dataset for light field deep learning," Institut für Industrielle Informationstechnik (IIIT), Karlsruhe, Germany, Tech. Rep. 1000125599, Sep. 2021, doi: 10.21227/y90t-xk47.
[47] M. Schambach and M. Heizmann, "A multispectral light field dataset and framework for light field deep learning," *IEEE Access*, vol. 8, pp. 193492–193502, 2020.
[48] P. Paudyal, F. Battisti, M. Sjöström, R. Olsson, and M. Carli, "Towards the perceptual quality evaluation of compressed light field images," *IEEE Trans. Broadcast.*, vol. 63, no. 3, pp. 507–522, Sep. 2017.
[49] R. R. Tamboli, B. Appina, S. Channappayya, and S. Jana, "Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display," *Signal Process., Image Commun.*, vol. 47, pp. 42–55, Sep. 2016.
[50] V. K. Adhikarla, M. Vinkler, D. Sumin, R. K. Mantiuk, K. Myszkowski, H.-P. Seidel, and P. Didyk, "Towards a quality metric for dense light fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 58–67.
[51] I. Viola and T. Ebrahimi, "VALID: Visual quality assessment for light field images dataset," in *Proc. 10th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2018, pp. 1–3.
[52] L. Shi, S. Zhao, W. Zhou, and Z. Chen, "Perceptual evaluation of light field image," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 41–45.
[53] L. Shan, P. An, D. Liu, and R. Ma, "Subjective evaluation of light field images for quality assessment database," in *International Forum on Digital TV and Wireless Multimedia Communications*. Cham, Switzerland: Springer, 2017, pp. 267–276.
[54] A. Zizien and K. Fliegel, "LFDD: Light field image dataset for performance evaluation of objective quality metrics," *Proc. SPIE*, vol. 11510, Aug. 2020, Art. no. 115102U.

[55] A. Gill, E. Zerman, C. Ozcinar, and A. Smolic, "A study on visual perception of light field content," in *Proc. Irish Mach. Vis. Image Process. Conf. (IMVIP)*, 2020, pp. 1–8.
[56] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," in *Proc. ACM SIGGRAPH Papers*, 2006, pp. 924–934.
[57] *The Stanford Multi-Camera Array*. Accessed: Dec. 5, 2020. [Online]. Available: http://graphics.stanford.edu/projects/array/
[58] D. G. Dansereau, "Light field toolbox v0.4—Deprecated, v0.5 available now," File Exchange, MATLAB Central, Version 1.1, Feb. 2015. [Online]. Available: http://mathworks.com
[59] P. A. Kara, A. Cserkaszky, S. Darukumalli, A. Barsi, and M. G. Martini, "On the edge of the seat: Reduced angular resolution of a light field cinema with fixed observer positions," in *Proc. 9th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2017, pp. 1–6.
[60] P. A. Kara, P. T. Kovacs, S. Vagharshakyan, M. G. Martini, A. Barsi, T. Balogh, A. Chuchvara, and A. Chehaibi, "The effect of light field reconstruction and angular resolution reduction on the quality of experience," in *Proc. 12th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, 2016, pp. 781–786.
[61] P. A. Kara, M. G. Martini, P. T. Kovacs, S. Imre, A. Barsi, K. Lackner, and T. Balogh, "Perceived quality of angular resolution for light field displays and the validy of subjective assessment," in *Proc. Int. Conf. 3D Imag. (IC3D)*, Dec. 2016, pp. 1–7.
[62] A. Sorkine-Hornung, C. Kim, H. Zimmer, Y. Pritch, and M. Gross, "Scene reconstruction from high spatio-angular resolution light fields," U.S. Patent 9 786 062, Oct. 10, 2017.
[63] B. O. Community, "Blender—A 3D modelling and rendering package," Blender Inst. B.V., Amsterdam, The Netherlands, Tech. Rep., 2016.
[64] *4D Light Field Dataset*. Accessed: Sep. 27, 2021. [Online]. Available: https://lightfield-analysis.uni-konstanz.de/
[65] P. Paudyal, R. Olsson, M. Sjöström, F. Battisti, and M. Carli, "SMART: A light field image quality dataset," in *Proc. 7th Int. Conf. Multimedia Syst.*, May 2016, pp. 1–6.
[66] M. Le Pendu, C. Guillemot, and A. Smolic, "A Fourier disparity layer representation for light fields," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5740–5753, Nov. 2019.
[67] *EyeLink 1000 Plus*. Accessed: Jul. 21, 2022. [Online]. Available: https://www.sr-research.com/eyelink-1000-plus/
[68] C. Meng, P. An, X. Huang, C. Yang, L. Shen, and B. Wang, "Objective quality assessment of lenslet light field image based on focus stack," *IEEE Trans. Multimedia*, vol. 24, pp. 3193–3207, 2022.
[69] H. PhiCong, S. Perry, E. Cheng, and X. HoangVan, "Objective quality assessment metrics for light field image based on textural features," *Electronics*, vol. 11, no. 5, p. 759, Mar. 2022.
[70] R. R. Tamboli, A. Cserkaszky, P. A. Kara, A. Barsi, and M. G. Martini, "Objective quality evaluation of an angularly-continuous light-field format," in *Proc. Int. Conf. 3D Immersion (IC3D)*, Dec. 2018, pp. 1–8.
[71] P. Paudyal, F. Battisti, and M. Carli, "Reduced reference quality assessment of light field images," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 152–165, Mar. 2019.
[72] L. Shan, P. An, C. Meng, X. Huang, C. Yang, and L. Shen, "A no-reference image quality assessment metric by multiple characteristics of light field images," *IEEE Access*, vol. 7, pp. 127217–127229, 2019.
[73] A. Cserkaszky, A. Barsi, P. A. Kara, and M. G. Martini, "To interpolate or not to interpolate: Subjective assessment of interpolation performance on a light field display," in *Proc. IEEE Int. Conf. Multimedia Expo. Workshops (ICMEW)*, Jul. 2017, pp. 55–60.
[74] P. A. Kara, R. R. Tamboli, A. Cserkaszky, M. G. Martini, A. Barsi, and L. Bokor, "The viewing conditions of light-field video for subjective quality assessment," in *Proc. Int. Conf. 3D Immersion (IC3D)*, Dec. 2018, pp. 1–8.
[75] P. A. Kara, R. R. Tamboli, E. Shafiee, M. G. Martini, A. Simon, and M. Guindy, "Beyond perceptual thresholds and personal preference: Towards novel research questions and methodologies of quality of experience studies on light field visualization," *Electronics*, vol. 11, no. 6, p. 953, Mar. 2022.
[76] P. A. Kara, A. Barsi, R. R. Tamboli, M. Guindy, M. G. Martini, T. Balogh, and A. Simon, "Recommendations on the viewing distance of light field displays," *Proc. SPIE*, vol. 11788, Jun. 2021, Art. no. 117880R.
[77] M. Martini, "IEEE standard on the quality assessment of light field imaging," in *Proc. IEEE SA*, Aug. 2021, pp. 20–55.
[78] S. Pezzulli, M. G. Martini, and N. Barman, "Estimation of quality scores from subjective tests-beyond subjects' MOS," *IEEE Trans. Multimedia*, vol. 23, pp. 2505–2519, 2021.
[79] J. Li and P. L. Callet, "Improving the discriminability of standard subjective quality assessment methods: A case study," in *Proc. 10th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2018, pp. 1–3.
[80] T. Hossfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!" in *Proc. 3rd Int. Workshop Quality Multimedia Exper.*, Sep. 2011, pp. 131–136.
[81] Z. Li and C. G. Bampis, "Recover subjective quality scores from noisy measurements," in *Proc. Data Compress. Conf. (DCC)*, Apr. 2017, pp. 52–61.
[82] L. Janowski and M. Pinson, "The accuracy of subjects in a quality experiment: A theoretical subject model," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2210–2224, Dec. 2015.

EDRIS SHAFIEE (Graduate Student Member, IEEE) received the B.Eng. degree in telecommunication engineering and the M.Sc. degree in computer vision and image analysis. He is currently pursuing the Ph.D. degree with Kingston University London. He has extensive industrial experience from his 20 years career in telecom sector. His research focuses on QoE of light field. He is a member and a Secretary of IEEE P.3333.1.4 Standards Association Working Group on the quality assessment of light field imaging. He is also a member of the Institution of Engineering and Technology (IET).

MARIA G. MARTINI (Senior Member, IEEE) received the Laurea degree (summa cum laude) in electronic engineering from the University of Perugia, Italy, in 1998, and the Ph.D. degree in electronics and computer science from the University of Bologna, Italy, in 2002. She is currently a Professor (Chair) at the Faculty of Science, Engineering, and Computing, Kingston University, London, U.K., where she also leads the Wireless and Multimedia Networking Research Group. She has led the KU Team in a number of national and international research projects, funded by the European Commission (OPTIMIX, CONCERTO, QoE-NET, and Qualinet), U.K. research councils (EPSRC, British Council, and Royal Society), U.K. Technology Strategy Board/Innovate U.K., and international industries. She has authored about 200 scientific articles, contributions to standardization groups (IEEE and ITU), and several patents on wireless video. Her research interests include QoE-driven wireless multimedia communications, decision theory, video quality assessment, and medical applications. She has chaired/organized a number of conferences and workshops. She is a part of international committees and expert groups, including the NetWorld2020 European Technology Platform Expert Advisory Group, the Video Quality Expert Group (VQEG), and the IEEE Multimedia Communications Technical Committee, where she has served as the Vice-Chair (2014–2016) and the Chair (2012–2014) of the 3-D Rendering, Processing, and Communications Interest Group. She was a key member of the QoE and multimedia streaming IG. She is an Expert Evaluator of the European Commission and EPSRC. She currently chairs the IEEE P.3333.1.4 Standards Association Working Group on the quality assessment of light field imaging. She was an Associate Editor of *IEEE Signal Processing Magazine* (2018–2020) and IEEE TRANSACTIONS ON MULTIMEDIA (2014–2018). She was also a Lead Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on QoE-Aware Wireless Multimedia Systems and the Guest Editor of the IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, IEEE MULTIMEDIA, and the *International Journal on Telemedicine and Applications*.